

AN INEQUALITY FOR EXPECTATION OF MEANS OF POSITIVE RANDOM VARIABLES

PAOLO GIBILISCO^{1*} and FRANK HANSEN²

ABSTRACT. Suppose that X, Y are positive random variable and m a numerical (commutative) mean. We prove that the inequality $E(m(X, Y)) \leq m(E(X), E(Y))$ holds if and only if the mean is generated by a concave function. With due changes we also prove that the same inequality holds for all operator means in the Kubo-Ando setting. The case of the harmonic mean was proved by C.R. Rao and B.L.S. Prakasa Rao.

1. INTRODUCTION AND PRELIMINARIES

Let x, y be positive real numbers. The arithmetic, geometric, harmonic, and logarithmic means are defined by

$$\begin{aligned} m_a(x, y) &= \frac{x+y}{2} & m_g(x, y) &= \sqrt{xy} \\ m_h(x, y) &= \frac{2}{x^{-1} + y^{-1}} & m_l(x, y) &= \frac{x-y}{\log x - \log y}. \end{aligned}$$

Suppose $X, Y: \Omega \rightarrow (0, +\infty)$ are positive random variables. Linearity of the expectation operator trivially implies

$$E(m_a(X, Y)) = m_a(E(X), E(Y)).$$

On the other hand the Cauchy-Schwartz inequality implies

$$E(m_g(X, Y)) \leq m_g(E(X), E(Y)).$$

Working on a result by Fisher on ancillary statistics Rao [11, 12] obtained the following proposition by an application of Hölder's inequality together with the harmonic-geometric mean inequality.

Proposition 1.1.

$$E(m_h(X, Y)) \leq m_h(E(X), E(Y)). \quad (1.1)$$

It is natural to ask about the generality of this result. For example, does it hold also for the logarithmic mean? To properly answer this question it is better to choose one of the many axiomatic approaches to the notion of a mean.

In Section 2 we recall the notion of *perspective* of a function, and in Section 3 we recall that a mean of pairs of positive numbers may be represented as the perspective of a certain representing function. In Section 4 we prove that inequality (1.1) holds for a mean m_f if and only if the representing function f is concave.

2010 *Mathematics Subject Classification.* Primary 26E60; Secondary 47A64, 60B20.

Key words and phrases. Numerical means, operator means, concavity, random matrices.

Once this is done it becomes natural to address the analog question in the non-commutative setting. A positive answer to the case of the matrix harmonic mean was given by Prakasa Rao in [10] and by C.R. Rao in [13]. But also in this case the inequality holds in a much wider generality. In Section 5 we recall the notion of non-commutative perspectives and some of their properties, while in Section 6 we describe the subclass of Kubo-Ando operator means. In Section 7 we show that inequality (1.1) holds true also in the non-commutative case. This follows from the fact that operator means are generated by operator monotone functions; indeed operator monotonicity of a function defined in the positive half-line implies operator concavity [6, Corollary 2.2]; rendering the non-commutative setting completely different from the commutative counter part.

In Section 8 we consider the random matrix case which, to some extent, encompasses the previous results.

2. PERSPECTIVE OF A FUNCTION: COMMUTATIVE CASE

Let $K \subseteq \mathbb{R}^n$ be a non-empty convex set, and let $g: K \rightarrow \mathbb{R}$ be a function. We consider the set

$$L = \{(x, t) \mid t > 0, t^{-1}x \in K\}.$$

Definition 2.1. The perspective \mathcal{P}_g of g is the function $\mathcal{P}_g: L \rightarrow \mathbb{R}$ defined by setting

$$\mathcal{P}_g(x, t) = tg(t^{-1}x) \quad (x, t) \in L.$$

The following classical result is well-known.

Proposition 2.2. *The perspective \mathcal{P}_g of a convex function g is convex.*

Example 2.3. Consider the convex function

$$g(x) = x \log x \quad x > 0$$

with limit $g(0) = 0$ and set $K = (0, \infty)$. Then the perspective is the relative entropy

$$\mathcal{P}_g(x, t) = x \log x - x \log t$$

for $x, t > 0$.

Notice that the perspective of a concave function is concave.

3. MEANS FOR POSITIVE NUMBERS

We use the notation $\mathbb{R}_+ = (0, +\infty)$.

Definition 3.1. A bivariate mean [9] is a function $m: \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$ such that

- (1) $m(x, x) = x$.
- (2) $m(x, y) = m(y, x)$.
- (3) $x < y \Rightarrow x < m(x, y) < y$.
- (4) $x < x'$ and $y < y' \Rightarrow m(x, y) < m(x', y')$.
- (5) m is continuous.
- (6) m is positively homogeneous; that is $m(tx, ty) = t \cdot m(x, y)$ for $t > 0$.

We use the notation \mathcal{M}_{num} for the set of bivariate means described above.

Definition 3.2. Let \mathcal{F}_{num} denote the class of functions $f: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ such that

- (1) f is continuous.
- (2) f is monotone increasing.
- (3) $f(1) = 1$.
- (4) $tf(t^{-1}) = f(t)$ for $t > 0$.

The following result is straightforward.

Proposition 3.3. *There is bijection between \mathcal{M}_{num} and \mathcal{F}_{num} given by the formulas*

$$m_f(x, y) = yf(y^{-1}x) \quad \text{and} \quad f_m(t) = m(1, t)$$

for positive numbers x, y and t .

3.1. Some examples of means. The functions in the table below are all concave, even operator concave.

TABLE 1.

Name	function	mean
arithmetic	$\frac{1+x}{2}$	$\frac{x+y}{2}$
WYD, $\beta \in (0, 1)$	$\frac{x^\beta + x^{1-\beta}}{2}$	$\frac{x^\beta y^{1-\beta} + x^{1-\beta} y^\beta}{2}$
geometric	\sqrt{x}	\sqrt{xy}
harmonic	$\frac{2x}{x+1}$	$\frac{2}{x^{-1} + y^{-1}}$
logarithmic	$\frac{x-1}{\log x}$	$\frac{x-y}{\log x - \log y}$

However, there exist non-concave functions in \mathcal{F}_{num} . Consider for example the function

$$g(x) = \frac{1}{4} \begin{cases} x+3 & 0 \leq x \leq 1, \\ 3x+1 & x \geq 1. \end{cases}$$

This piece-wise affine function is convex and belongs to \mathcal{F}_{num} .

4. THE MAIN RESULT: COMMUTATIVE CASE

Theorem 4.1. *Take a function $f \in \mathcal{F}_{num}$. The inequality*

$$E(m_f(X, Y)) \leq m_f(E(X), E(Y)) \tag{4.1}$$

holds for arbitrary positive random variables X and Y if and only if f is concave.

Proof. Suppose inequality (4.1) holds for a function f . Take $\Omega = \{1, 2\}$ as state space with probabilities p and $1 - p$, and let Y be the constant function 1. We set $X(1) = x_1$ and $X(2) = x_2$ for given $x_1, x_2 > 0$. We then have $E(Y) = 1$ and thus

$$m_f(E(X), E(Y)) = E(Y)f\left(\frac{E(X)}{E(Y)}\right) = f(px_1 + (1 - p)x_2).$$

We also have

$$m_f(X, Y)(1) = Y(1)f\left(\frac{X(1)}{Y(1)}\right) = f(x_1)$$

and

$$m_f(X, Y)(2) = Y(2)f\left(\frac{X(2)}{Y(2)}\right) = f(x_2).$$

Therefore

$$\begin{aligned} pf(x_1) + (1 - p)f(x_2) &= E(m_f(X, Y)) \leq m_f(E(X), E(Y)) \\ &= f(px_1 + (1 - p)x_2) \end{aligned}$$

implying that f is concave.

Suppose on the other hand that f is concave and consider two positive random variables X and Y . We only have to prove the theorem under the assumption that X and Y are simple random variables (finite linear combinations of indicator functions). The general case then follows since any positive random variable is a pointwise increasing limit of simple random variables. The (different) values of X are denoted by x_1, \dots, x_n with associated (marginal or unconditional) probabilities p_1, \dots, p_n . The (different) values of Y are denoted by y_1, \dots, y_m with associated (marginal or unconditional) probabilities q_1, \dots, q_m .

The stochastic variable $m_f(X, Y)$ takes the values $m_f(x_i, y_j)$ with probabilities $P(X = x_i, Y = y_j)$ for $i = 1, \dots, n$ and $j = 1, \dots, m$ (possibly counted with multiplicity). The mean m_f is the perspective of f and thus concave by Proposition 2.2. We may therefore apply Jensen's inequality and obtain

$$\begin{aligned} E(m_f(X, Y)) &= \sum_{i=1}^n \sum_{j=1}^m P(X = x_i, Y = y_j) m_f(x_i, y_j) \\ &\leq m_f\left(\sum_{i=1}^n \sum_{j=1}^m P(X = x_i, Y = y_j)(x_i, y_j)\right) \\ &= m_f\left(\sum_{i=1}^n \sum_{j=1}^m P(X = x_i, Y = y_j)x_i, \sum_{j=1}^m \sum_{i=1}^n P(X = x_i, Y = y_j)y_j\right), \end{aligned}$$

where we interchanged the summations in the second argument of m_f . Since the sums of the joint probabilities

$$\sum_{j=1}^m P(X = x_i, Y = y_j) = p_i \quad \text{and} \quad \sum_{i=1}^n P(X = x_i, Y = y_j) = q_j$$

we obtain

$$E(m_f(X, Y)) \leq m_f \left(\sum_{i=1}^n p_i x_i, \sum_{j=1}^m q_j y_j \right) = m_f(E(X), E(Y)),$$

which is the desired inequality (4.1). **QED**

5. NON-COMMUTATIVE PERSPECTIVE

For the basic results of this section we refer to [1, 3, 2]. Let f be a function defined in the open positive half-line. In Section 2 we recalled the perspective of f as the function of two variables $\mathcal{P}_f(t, s) = sf(s^{-1}t)$, where $t, s > 0$. Depending on the application, we may also consider the function $(t, s) \rightarrow \mathcal{P}_f(s, t)$ and denote this as the perspective of f .

If A and B are commuting positive definite matrices, then the matrix $\mathcal{P}_f(A, B)$ is well-defined by the functional calculus, and it coincides with $Bf(B^{-1}A)$. However, even if A and B do not commute one may, by choosing an appropriate ordering, define the perspective.

Definition 5.1. Let f be a function defined in the open positive half-line. The (non-commutative) perspective \mathcal{P}_f of f is then defined by setting

$$\mathcal{P}_f(A, B) = A^{1/2} f(A^{-1/2} B A^{-1/2}) A^{1/2}$$

for positive definite operators A and B .

For the following basic result confer [1, Theorem 2.2], [2, Theorem 1.1] and [3, Theorem 2.2].

Theorem 5.2. *The (non-commutative) perspective \mathcal{P}_f is convex if and only if f is operator convex.*

Let $f: (0, \infty) \rightarrow \mathbf{R}$ be a convex function. Since the perspective \mathcal{P}_f is both convex and positively homogenous we obtain the inequality

$$\mathcal{P}_f \left(\sum_{i=1}^n \lambda_i x_i, \sum_{i=1}^n \lambda_i y_i \right) \leq \sum_{i=1}^n \lambda_i \mathcal{P}_f(x_i, y_i)$$

for tuples (x_1, \dots, x_n) and (y_1, \dots, y_n) of positive numbers and positive numbers $\lambda_1, \dots, \lambda_n$. This entails, by setting all $\lambda_i = 1$, the inequality

$$\mathcal{P}_f(\text{Tr } A, \text{Tr } B) \leq \text{Tr } \mathcal{P}_f(A, B)$$

for commuting positive definite matrices A and B .

The transformer inequality for the non-commutative perspective of an operator convex function is essentially proved in [5, Theorem 2.2]. Since the perspective of an operator convex function is a convex regular operator map the statement also follows from [7, Lemma 2.1].

Proposition 5.3 (the transformer inequality). *Let $f: (0, \infty) \rightarrow \mathbf{R}$ be an operator convex function. The non-commutative perspective \mathcal{P}_f satisfies the inequality*

$$\mathcal{P}_f(C^* A C, C^* B C) \leq C^* \mathcal{P}_f(A, B) C$$

for every contraction C and positive definite operators A and B .

Notice that we by homogeneity obtain

$$\mathcal{P}_f(C^*AC, C^*BC) \leq C^*\mathcal{P}_f(A, B)C$$

for any operator C . In particular, if C is invertible we then have

$$\mathcal{P}_f(A, B) \leq (C^*)^{-1}\mathcal{P}_f(C^*AC, C^*BC)C^{-1} \leq \mathcal{P}_f(A, B),$$

hence there is equality and thus

$$C^*\mathcal{P}_f(A, B)C = \mathcal{P}_f(C^*AC, C^*BC). \quad (5.1)$$

Proposition 5.4. *Let \mathcal{P}_f be the non-commutative perspective of an operator convex function $f: (0, \infty) \rightarrow \mathbb{R}$ and let c_1, \dots, c_n be operators on a Hilbert space \mathcal{H} such that $c_1^*c_1 + \dots + c_n^*c_n = 1$. Then*

$$\mathcal{P}_f\left(\sum_{i=1}^n c_i^* A_i c_i, \sum_{i=1}^n c_i^* B_i c_i\right) \leq \sum_{i=1}^n c_i^* \mathcal{P}_f(A_i, B_i) c_i$$

for positive definite operators A_1, \dots, A_n and B_1, \dots, B_n acting on \mathcal{H} .

Proof. The perspective \mathcal{P}_f is a convex regular operator map of two variables [5, 2, 7]. The statement thus follows from Jensen's inequality for convex regular operator maps [7, Theorem 2.2]. **QED**

6. OPERATOR MEANS IN THE SENSE OF KUBO-ANDO

The celebrated Kubo-Ando theory of matrix means [8, 9, 4] may today be considered as part of the theory of perspectives of positive operator concave functions. This setting is simpler than the general theory of perspectives since a positive operator concave function necessarily is increasing, while a positive operator convex function may not necessarily be monotonic.

Definition 6.1. A bivariate *mean* for pairs of positive operators is a function

$$(A, B) \rightarrow m(A, B)$$

defined in and with values in positive definite operators on a Hilbert space and satisfying, mutatis mutandis, conditions (1) to (5) in Definition 3.1. In addition the *transformer inequality*

$$C^*m(A, B)C \leq m(C^*AC, C^*BC)$$

holds for positive definite A, B and arbitrary C .

Notice that the transformer inequality replaces (6) in Definition 3.1. We denote by \mathcal{M}_{op} the set of matrix means.

Example 6.2. The arithmetic, geometric and harmonic (matrix) means are defined, respectively, by setting

$$\begin{aligned} A \nabla B &= \frac{1}{2}(A + B) \\ A \# B &= A^{1/2} (A^{-1/2} B A^{-1/2})^{1/2} A^{1/2} \\ A ! B &= 2(A^{-1} + B^{-1})^{-1}. \end{aligned}$$

We recall that a function $f: (0, \infty) \rightarrow \mathbb{R}$ is said to be *operator monotone (increasing)* if

$$A \leq B \Rightarrow f(A) \leq f(B)$$

for positive definite operators on an arbitrary Hilbert space. An operator monotone function f is said to be *symmetric* if $f(t) = tf(t^{-1})$ for $t > 0$ and *normalized* if $f(1) = 1$.

Definition 6.3. \mathcal{F}_{op} is the class of functions $f: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ such that

- (1) f is operator monotone increasing,
- (2) $tf(t^{-1}) = f(t) \quad t > 0$,
- (3) $f(1) = 1$.

The fundamental result, due to Kubo and Ando, is the following.

Theorem 6.4. *There is bijection between \mathcal{M}_{op} and \mathcal{F}_{op} given by the formula*

$$m_f(A, B) = A^{1/2} f(A^{-1/2} B A^{-1/2}) A^{1/2}.$$

Remark 6.5. All the function in \mathcal{F}_{op} are (operator) concave making the operator case quite different from the numerical one.

If ρ is a density matrix and A is self-adjoint then the expectation of A in the state ρ is defined by setting $E_\rho(A) = \text{Tr}(\rho A)$.

7. THE MAIN RESULT: NONCOMMUTATIVE CASE

Theorem 7.1. *Take $f \in \mathcal{F}_{op}$. Then*

$$E_\rho(m_f(A, B)) \leq m_f(E_\rho(A), E_\rho(B)), \quad (7.1)$$

Proof. Consider a spectral resolution

$$\rho = \sum_{i=1}^n \lambda_i e_i$$

of the density matrix ρ in terms of one-dimensional orthogonal eigenprojections e_1, \dots, e_n with corresponding eigenvalues $\lambda_1, \dots, \lambda_n$ counted with multiplicity. By setting $c_i = \lambda_i^{1/2} e_i$ for $i = 1, \dots, n$ we obtain

$$E_\rho(A) = \text{Tr} \rho A = \text{Tr} \sum_{i=1}^n c_i^* A c_i$$

for any operator A . By using the transformer inequality we obtain

$$\begin{aligned} E_\rho(m_f(A, B)) &= \text{Tr} \sum_{i=1}^n c_i^* m_f(A, B) c_i \\ &\leq \text{Tr} m_f\left(\sum_{i=1}^n c_i^* A c_i, \sum_{i=1}^n c_i^* B c_i\right) \\ &\leq m_f\left(\text{Tr} \sum_{i=1}^n c_i^* A c_i, \text{Tr} \sum_{i=1}^n c_i^* B c_i\right) \\ &= m_f(E_\rho(A), E_\rho(B)), \end{aligned}$$

where we in the second inequality used that the operators

$$\sum_{i=1}^n c_i^* A c_i \quad \text{and} \quad \sum_{i=1}^n c_i^* B c_i$$

are commuting. **QED**

8. THE RANDOM MATRIX CASE

Let (Ω, \mathcal{F}, P) be a probability space. A map $X: \Omega \rightarrow M_n$ is called a random matrix. We may write

$$X = (X_{i,j})_{i,j=1}^n: \Omega \rightarrow M_n$$

and say that X is a positive definite random matrix if

$$X(\omega) = (X_{i,j}(\omega))_{i,j=1}^n$$

is positive definite for P -almost all $\omega \in \Omega$. We may readily consider other types of definiteness for random matrices.

Definition 8.1. A positive semi-definite random matrix $\rho: \Omega \rightarrow M_n$ is called a random density matrix if $\text{Tr } \rho = 1$ for P -almost all $\omega \in \Omega$.

Let X and ρ be random matrices on the probability space (Ω, \mathcal{F}, P) and suppose that ρ is a random density matrix. We introduce the pointwise expectation $E_\rho(X)$ by setting

$$(E_\rho X)(\omega) = \text{Tr } \rho(\omega) X(\omega) \quad \omega \in \Omega.$$

The pointwise expectation $E_\rho(X)$ is a random variable with mean

$$E(E_\rho(X)) = \int_\Omega \text{Tr } \rho(\omega) X(\omega) dP(\omega).$$

If ρ is a constant density matrix then

$$E(E_\rho(X)) = \text{Tr } \rho \int_\Omega X(\omega) dP(\omega) = \text{Tr } \rho E(X) = E_\rho(E(X)),$$

where $E(X)$ is the constant matrix with entries

$$E(X)_{i,j} = \int_\Omega X_{i,j}(\omega) dP(\omega) \quad i, j = 1, \dots, n.$$

Theorem 8.2. Let X and Y be positive definite random matrices on a probability space (Ω, \mathcal{F}, P) . For $f \in \mathcal{F}_{op}$ we obtain the inequality

$$E E_\rho(m_f(X, Y)) \leq m_f(E E_\rho(X), E E_\rho(Y))$$

for each random density matrix ρ on (Ω, \mathcal{F}, P) .

Proof. The matrices $X(\omega)$, $Y(\omega)$ and $\rho(\omega)$ are positive definite and $\rho(\omega)$ has unit trace for almost all $\omega \in \Omega$. The inequality between random variables

$$E_{\rho(\omega)}(m_f(X(\omega), Y(\omega))) \leq m_f(E_{\rho(\omega)}(X(\omega)), E_{\rho(\omega)}(Y(\omega)))$$

is therefore valid by our non-commutative inequality in Theorem 7.1. In particular, by taking the mean on both sides, we obtain

$$\begin{aligned} \mathbb{E} \mathbb{E}_\rho(m_f(X, Y)) &\leq \mathbb{E}(m_f(\mathbb{E}_\rho(X), \mathbb{E}_\rho(Y))) \\ &\leq m_f(\mathbb{E} \mathbb{E}_\rho(X), \mathbb{E} \mathbb{E}_\rho(Y)), \end{aligned}$$

where we used, in the last inequality, the commutative inequality in Theorem 4.1. **QED**

Notice that Theorem 8.2 reduces to the non-commutative inequality when Ω is a one point space, and to the commutative inequality when $n = 1$. If ρ is a constant matrix then the order of \mathbb{E} and \mathbb{E}_ρ in the inequality may be reversed.

Acknowledgments. It is a pleasure for the first author to thank Fumio Hiai for discussions and hints on the subject. The second author acknowledges support from the Japanese government Grant-in-Aid for scientific research 26400104.

REFERENCES

1. E.G. Effros. *A matrix convexity approach to some celebrated quantum inequalities*. Proc. Natl. Acad. Sci. USA, **106** (2009), no. 4, 1006–1008.
2. E.G. Effros and F. Hansen. *Non-commutative perspective*. Ann. Funct. Anal., **5**, (2014), no. 2, 74–79.
3. A. Ebadian, I. Nikoufar and M.E. Gordji. *Perspective of matrix convex functions*. Proc. Natl. Acad. Sci. USA, **108** (2011), no. 18, 7313 – 7314.
4. P. Gibilisco, F. Hansen and T. Isola. *On a correspondence between regular and non-regular operator monotone functions*. Lin. Alg. Appl., **430** (2009), no. 8–9, 2225–2232.
5. F. Hansen. *Means and concave products of positive semidefinite matrices*. Math Ann., **264** (1983), 119 – 128.
6. F. Hansen. *The fast track to Löwner’s theorem*. Linear Algebra Appl., **438** (2013), 4557 – 4571.
7. F. Hansen. *Perspectives and completely positive maps*. arXiv:1601.06723v4 [math-ph] (2016), to appear in Ann. Funct. Anal.
8. F. Kubo and T. Ando. *Means of positive linear operators*. Math. Ann. **246** (1979/80), no. 3, 205–224.
9. D. Petz and R. Temesi. *Means of positive numbers*. SIAM J. Matrix Anal. Appl., **27** (2005), 712–720.
10. B.L.S. Prakasa Rao. *An inequality for the expectation of harmonic mean of random matrices*. Technical Report, Indian Statistical Institute, Delhi, (1998).
11. C.R. Rao. *R.A. Fisher: the founder of modern statistics*. Statistical Science, **7** (1992), 34–48.
12. C.R. Rao. *Seven inequalities in statistical estimation theory*. Student, **1** (1996), no. 3, 149–158.
13. C.R. Rao. *Statistical proofs of some matrix inequalities*. Lin. Alg. Appl., **321** (2000), 307 – 320.

¹DEPARTMENT OF ECONOMICS AND FINANCE, UNIVERSITY OF ROME “TOR VERGATA”, VIA COLUMBIA 2, ROME 00133, ITALY.

E-mail address: gibilisco@volterra.uniroma2.it

²INSTITUTE FOR EXCELLENCE IN HIGHER EDUCATION, TOHOKU UNIVERSITY, SENDAI, JAPAN.

E-mail address: frank.hansen@m.tohoku.ac.jp